
Implementing screens within screens to create telepresence in reality TV

Daniel Klug * ¹, Axel Schmidt * † ²

¹ Institute for Media Studies, University of Basel – Suisse

² Institute for the German Language – Allemagne

Following Gumbrecht's (2004) draft of a theory of presence the *notion of presence* aims at spatial relations: Something is *present* if it is tangible or 'in actual range' (Sch'utz/Luckmann 1979). Concerning the *social world*, Goffman (1971) understands *co-presence* in a similar way, as the chance to get in direct contact with each other. For this, the possibility of *mutually perceiving each other* is a precondition and is in turn made possible by a *shared location*. Goffman's concept of co-presence is bound to a *shared here and now* and thus is based on a *notion of situation* with *place* as its central dimension. *Interaction* as a basic mode of human exchange builds on situations and the possibility of mutual perception allowing so called *perceived perceptions* in particular (Hausendorf 2003) as the main mechanism of focused encounters.

Media studies, however, proved that both the notion of *situation* as place and *interaction* as co-presence need to be *extended* as mediated contact creates new forms of situations and interactions. Especially Meyrowitz (1990) shows that the notion of situation has to be detached from the dimension of place and instead be redefined in terms of *mediated contact*. Therefore, the question "who can hear me, who can see me?" is the basis for individual orientation and should guide the analysis of (inter)action. Although media enlarges the scope of perception in this way, it is not just a simple extension, but a *transformation process* in which the *mediation itself* plays a crucial role.

A. Extending and shifting presence in television

Transferred to *screen based media*, in our case *television*, presence gets extended to a notion of *telepresence* (i.a. Rasmussen 2000) or *presence at a distance* (Meyrowitz 1990) where others are not spatially present but *perceivable* in great detail (with different effects depending on the form of communication, e.g. unidirectional mass media vs. bi-directional interpersonal media). This requires the ability of *audiovisual media* to bridge distances between separate locations by conveying images and sounds which are *framed as portrayals of real events*. Goffmann (1979) defines this as one function of photographic images which he calls "instantiation" (20) and by which the existence of an event is verified by framing an image as 'instantiation'. This is enabled by corresponding techniques which cover real

*Intervenant

†Auteur correspondant: axel.schmidt@ids-mannheim.de

events ‘*similar to perception*’ (Sachs-Hombach 2003) and preserve them for later usage and/or broadcast them in the very moment with the latter being a *live transmission* in terms of a special form of presence. If screen based communication becomes *interactive* additional levels of telepresence emerge either on an *interpersonal level* (e.g. Skype), on a *mass media level* (e.g. video interviews within talk shows) or an ‘*interaction media*’ level (Thiedeke 2006) (e.g. YouTube as ‘interaction of audiovisual artefacts’).

However, presence not only gets extended by screen based media but also shifted:

1. *Time shifts*: Presented (inter-)actions do not need to take place at the very moment but might be recorded in the past and are re-actualized later thereby bridging a temporal gap.
2. *Shifts in the framing of interactions*: Situations presented within television are not just situations but ‘*situations of representation*’ (i.a. Keppler 1994). That means interaction itself is depicted for an audience, thus understandable as ‘*interaction staging*’. Thereby not only the depicted event itself is *heightened* in status by being selected and exposed, but also the behavior of the depicted persons is seen as tending towards *performance* due to the camera coverage.
3. *Shifts on the level of interactional coordination* as basic forms are altered through mediation, e.g. the awareness of glances or postural shifts as in the case of screen mediated interpersonal interaction (Heath/Luff 1993).

This paper focuses on so-called *reality TV* in which staged events with ‘real people’ serve to maintain a ‘claim to the real’ (Jermyn/Holmes 2004). Therefore, it is dealing with a mixture of professional contexts (TV) and (apparently) private ones (depicted everyday situations/people) shown for entertainment purposes within the mass medium of television. The paper discusses a *special practice* which makes events/persons ‘*telepresent*’ within the formats, i.e. focusing on ‘*telepresented telepresence*’ or ‘*screens within screens*’. A lot of reality TV-shows *use (additional) screens* to present participant behavior either recorded in the past or live. For example, this is found in shows like Big Brother when television screens are used in the studio setting to depict simultaneous events in the container house that are (seemingly) being watched by the hosts; Or in *makeover* shows (Heller 2007) like The Swan when video footage of past events is watched together to confront participants with their former behavior or appearance, e.g. to create emotional backlashes and dramatizations based on previous failures.

Methodologically, analyzing the details of how screens are implemented within reality TV formats is done using multimodal video analysis (Norris 2004) and television analysis (i.a. Mikos 2003). In addition, to improve the understanding of how and why screens are embedded within television programs a *typology* is developed.

To do so, the following analytical *levels* are crucial:

Level of depicted events: How screens are embedded interactionally within the ongoing course of activities in turn producing screen-based interactions on screen (e.g. people in TV shows using screens/video); *Level of depiction*: Why and with what intention/function screens are implemented within the formats; *Interplay of the two*

levels: How changes within interactions relate to functions designed for viewers; are there systematic relations between an interactional and a depicted level?

In addition, amongst others the following *analytical dimensions* are of interest:

- Who produced the embedded videos?
- By whom are they introduced and how are they fitted into the ongoing activities?
- What is the recipient setting like ('dipositif', spatial arrangement etc.)?
- What does the camera show in which order (screen, participants watching/interacting, both, the whole setting, switches between several views etc.)?
- What do the participants do while watching and for what purposes?
- Who does (not) benefit from the exposure of the video(s)?
- Is the video content part of the running show or the result of external creations (textually/spatially and/or temporally)?
- Are videos played once or several times? With(out) interruption(s)?
- Is the video an expectable or exceptional part of the show (also for the participants)?
- Were the people on the second screen aware of the recording?
- Do the participants on the first screen interact with persons showed on the second screen or do they just watch?
- How does the show as a fabricated narrative build on the integrated video(s) in the further course of the show?

B. Examples of screen in screen usage in reality TV

Following the above dimensions reality TV shows of various genres can be identified using forms of additional screens for different narrative and dramatic reasons. The most obvious televisual usage of second screens can be found in reality-based TV shows like MTV's *Ridiculousness* or the classic *America's Funniest Home Videos* that are showing various clips to a present guest and/or audience. Though the short clips are shown as single screen content to the television audience it is evident that the show's guest did see them too as they are later commenting on the clips. This concept imitates the general perception situation of television content and transfers the content based communication into the entertaining concept of the show.

The most practical usage might the screen-based video call interaction in TV shows like the MTV shows *Catfish – The TV Show* or *Say It In Song* in which the hosts of the show are seen talking to the constants who are e.g. seeking help to find other people

or getting into contact with them. Yet, this already illustrates the bridging of the spatial gap for example to include otherwise inaccessible visualizations of the contestants' emotions through facial expressions. Said shows generally use these screen based interactions in the first third of the narrative to construct a climax before a first confrontation.

A mixture of these two usages can be found for example in the reality show *Wife Swap* (and national versions) where the contestants prerecord messages to instruct each other during the swapping experiment. The contestant then is shown watching the video and is constrained to verbally interact with the recorded message creating a 'person-to-screen(-person)' parasocial interaction. This screen/video usage happens at the beginning of an episode (instructions) and in the last third (telling about conflicts in the new family situations). It is mainly used to create further tension between the generally opposing contestants and lifestyles.

Another common usage of additional screens is to provide an insight into simultaneous show actions to a spatially remote live audience in the studio or to the hosts, mostly in competitive shows like *I'm a celebrity... Get me out of here!* or *Big Brother* (also national versions). While the first screen presents the live show, the hosts often use the second screen to comment on selected events or contestant behavior to pre-frame the following show narrative. Casting shows like *American Idol* or *America's Got Talent* use TV screens mostly in the first couple episodes to show (some of) the contestants' performances in front of the jury to people (oftentimes family members) in the waiting areas who are sometimes commenting on it.

A distinct usage of screens is seen in coaching shows like *Supernanny* (e.g. US, UK, Germany, Austria) where second screens are used to show prerecorded (documentary) audiovisual material to the contestants in order to confront them with their (mis-)behavior. After the first depiction of the deviant family situation, the videos are watched together by the parents and the coach who uses the video footage as evidence to intervene into the parents educational methods which is framed as a coaching of the family. This screen usage aims at internal confrontations of the parents, emotional confessions, and at the verbalization of help by the parents.

Nowadays the highest screen in screen usage can be found in MTV dating shows. Following different dating concepts, most of the shows are using second screens to either create information gaps between the male and female contestants or to give the television audience additional insights on the televisual events. A significant aspect is, that the second screen, hence the show narrative's additional information, in general is not shown but made inferable through the contestants' receptive actions. For example, in *Room Raiders* one person examines the rooms of three unknown persons to choose a date while these three people watch him/her going through their belongings. Yet, the watchers are filmed from a front perspective in which the camera substitutes the actual second screen. Therefore, they constantly interact with the camera (that generates the first screen) by commenting on the actions of the filmed person. Similar concepts can be found in (former) MTV surveillance shows like *Parental Control*, *Exposed*, or *Damage Control* in which parents prank their kids by making them believe that they are absent while the kids actions are secretly filmed and commented on by the watching parents.

The entertaining use of additional screens to observe some contestants to provide other absent contestants with exclusive audiovisual information mostly combines relationship and gaming aspects of reality TV. For example, in the MTV dating show *The X Effect* two people are put in a hotel room while both their partners (who are exes) spend time in a different room. The other two are given the possibility to watch their partners' actions on a surveillance screen which before that shows them a map of the room and the location of their partners. In this way, dating reality exploits possible suspicions and the fear of cheating by only providing pre-selected or marginal information about simultaneous but spatially inaccessible events.

Reality shows like *The X Effect* therefore are good examples to analyze how screen in screen based audiovisual content provides narrative bases for said shows that are centered on alternating interactions of contestants, hosts, and audiences. The usages of screen content can be differed by the involved parties and their relation to the shown/watched televisual events, e.g.: communicating and commenting by third person parties like hosts or guests; observing and confronting by second person parties like coaches or experts; and instructing and reporting by first person parties like the contestants themselves.

C. Notes on a possible typology

I. Ways of depiction

1. *Border case: montage/editing*

Filmic/cinematic elements like switching between different time frames (flashback, pre-view etc., e.g. a challenge by a candidate in *Big Brother*) are *not considered* as cases in point because they are just time shifts on a narrative level. That means that depictions generated in this way are not framed as 'watching something together' cause only the TV viewers can see these events.

2. *Simple case: replacement (features) and the staging of 'watching together'*

Replacements consecutively present two different events on one screen as the actual screen presenting the show (situation 1) is temporally *replaced* by a screen that shows a feature (e.g. a past event = embedded situations 2-n). This is often done in factual formats like news or talk shows. The succession of images/events and its implementation (via a second screen) suggest that people within the ongoing situation (situation 1, e.g. talk show guests, host and live audience) are simultaneously *watching the feature together on a screen* (situation 2) provided within the actual show (situation 1) although this watching of the screen itself is not presented. This practice can be understood as providing '*second screens*' (= situation 2 = feature) within a '*first screen*' (= situation 1 = live event) yet not depicting the whole situation of watching and rather *replacing* one screen temporally while implying an invisible process of watching in the meantime. This gets evident after returning back the show as host and guest(s) usually referring to the feature which was shown what indicates a symmetrical state of knowledge. So, the *first screen serves as a basis* showing an ongoing event (in this case a show) which continues even if it is not made visible and to which the depiction thus can return. Those screens we call '*anchoring screens*'. However, in the next step *images of the situation of watching the second screen*

are provided on the anchoring screen which leads to different realizations:

3. *Editing: perception front / event on second screen*

Watching the ‘second screen’ can be shown by *successively editing* a) front views of the *watchers* and b) *what is being watched on the second screen* in a parallel montage. This focuses on reactions of the watchers (as their faces are captured) and the unfolding of the watched event (as it is shown in a ‘replacement mode’, see 2. *Simple case*). Here the sound editing of the anchoring screen (e.g. talk about the watched events) is crucial as it generally can be heard with the images of the second screen.

4. *Whole situation: Perception back + screen (events)*

Furthermore, it can be realized by depicting the *whole situation of reception*. This usually requires a back view to cover both the *watchers* (from behind in the foreground) and the *watched event on the second screen* (in the background). Here especially immediate bodily reactions of the watchers are captured as well as the events on the second screen itself. The peculiarity is that the watching and the watched get synchronized in one screen and the relevant information on the actual second screen are made visible within the anchoring screen.

5. *Realization by screen-technical means (split screen, picture in picture)*

In addition to the above implementations, this can also be based on technical means of visual filmmaking. The anchoring screen can be technically manipulated to present a second screen, for example as split screen or picture-in-picture mode. In this way, a simultaneous view of two (or more) spatially separated events is provided in two (or more) images (or: subordinated screens) within one major screen connected by the (claimed and/or presumed) relation of watchers and watched events (e.g. first screen = failure of a candidate in the past, second screen = candidate’s reactions while watching his previous failure).

As shown in the last case, implementing second screens is based on different *levels of production*:

II. Levels of production

The use of a second screen concerns different *levels of realization* depending on the *way of depiction* (see above):

editing/montage: anchoring screen (watching) and (events on) second screen (watched) are successively arranged by (live) editing (post production 1); *shooting*: watching and (events on) second screen (watched) are shown as one situation within the anchoring screen (back perspective + event) (production); *screen realization/ image mixing*: anchoring screen (watching) and (events on) second screen (watched) are shown as single situations but simultaneously on one screen by means of image mixing (post production 2).

D. Possible functions

Concerning the interplay of interaction and depiction level, the implementation of video in reality TV provides multiple functions as it is a medium to regulate ongoing events by bringing in other (previous) events that are framed as real. This is an effective way to shape ongoing events interactionally. Furthermore, screens in screens are used to authenticate entire shows. For example, a live transmission to an artificial setting (like the Big Brother container) within a live show is designed to demonstrate the (parallel) existence of this 'world' to which the live show refers and thereby builds upon.

Also, video (pre-)recordings function to stretch the actual narrative time by duplicating events or including preceding events leading to alternations in the screen presence of contestants (see B.).

References

Goffman, Erving (1971): *Verhalten in sozialen Situationen. Strukturen und Regeln der Interaktion im 'öffentlichen Raum*. G'utersloh: Bertelsmann.

Goffman, Erving (1979): *Gender advertisements*. New York: Harper Row.

Gumbrecht, Hans Ulrich (2004): *Diesseits der Hermeneutik. die Produktion von Pr'asenz*. Frankfurt/M.: Suhrkamp.

Hausendorf, Heiko (2003): "Deixis and speech situation revisited: the mechanism of perceived perception". In: Lenz, Friedrich (Hg.): *Deictic conceptualisation of space, time and person*. Amsterdam: John Benjamins, pp. 249–269.

Heath, Christian/Luff, Paul (1993): "Disembodied Conduct: Interactional Asymmetries in Video-Mediated Communication". In: Button, Graham (Ed.): *Technology in working order - studies of work, interaction, and technology*. London: Routledge, pp. 35–54.

Heller, Dana (2007): *Makeover television: realities remodeled*. London: Tauris.

Holmes, Sue/Jermyn, Deborah (2004): "Introduction". In: Holmes, Sue/Jermyn, Deborah (Ed.): *Understanding reality television*. London: Routledge, pp. 1–32.

Keppler, Angela (1994): *Wirklicher als die Wirklichkeit? Das neue Prinzip der Fernsehunterhaltung*. Frankfurt/M: Fischer.

Meyrowitz, Joshua (1990): "Redefining the Situation: Extending Dramaturgy into a theory of social change and media effects". In: Riggins, Stephen H. (Ed.): *Beyond Goffman: studies on communication, institution, and social interaction*. Berlin [u.a.]: Mouton de Gruyter, pp. 65–97.

Mikos, Lothar (2003): *Film- und Fernsehanalyse*. UVK: Konstanz.

Norris, Sigrid (2004): *Analyzing multimodal interaction*. London: Routledge.

Rasmussen, Tove A. (2000): "Construction of Authenticity. Television talk as situated interaction". In: Hallenberger, Gerd/Schanze, Helmut (Ed.): *Live is Life. Mediale Inszenierungen des Authentischen*. Baden-Baden: Nomos, pp. 47–67.

Sachs-Hombach, Klaus (2003): „Vom Bild zum Film. Zur begrifflichen Analyse wahrnehmungsnaher Kommunikation“. In: Ehrenspeck, Yvonne/Schäffer, Burkhard (Ed.): *Film- und Fotoanalyse in den Erziehungswissenschaften*. Opladen: Leske+Budrich, pp. 121–134.

Schütz, Alfred/Luckmann, Thomas (1979): *Strukturen der Lebenswelt*. Band 1. Frankfurt/M: Suhrkamp.

Thiedeke, Udo (2006): „Interaktionsmedien“. In: Tsvasman, Leon R. (Ed.): *Das große Lexikon Medien und Kommunikation: Kompendium interdisziplinärer Konzepte*. Würzburg: Ergon, pp. 160–163.

Mots-Clés: reality television, telepresence, surveillance, depicted events, screen, based interaction, multimodal video analysis